1 [Counsel on signature page]

2

3

4

5

6

7 **UNITED STATES DISTRICT COURT**
NORTHERN DISTRICT OF CALIFORNIA
SAN FRANCISCO DIVISION

8

9 Stewart O'Nan, et al.   Case No. 3:24-cv-01451-CRB
                                                Case No. 3:24-cv-02653-CRB
10   *Individual and Representative Plaintiffs,*

11   vs.   **JOINT CASE MANAGEMENT**
                                                STATEMENT AND RULE 26(f) REPORT
12 Databricks, Inc., et al.,
                                                Hearing Date: June 21, 2024
13   *Defendants.*   Time: 8:30 AM
                                                Place: Courtroom F, 15th Floor
14 ─────────────────────────────────   Judge: Hon. Charles R. Breyer

15 Rebecca Makkai, et al.,

16   *Individual and Representative Plaintiffs,*

17   vs.

18 Databricks, Inc., et al.,

19   *Defendants.*

20

21

22

23

24

25

26

27

28
                                                Case No. 3:24-cv-01451-CRB
                                                Case No. 3:24-cv-02653-CRB

1   Pursuant to Federal Rule of Civil Procedure 26(f)(2), Civil Local Rule 16-9(a), and the

2   Court's Case Management Orders Dkt. 8[1] (*O'Nan* docket) and Dkt. 4 (*Makkai* docket)[2], counsel for

3   Plaintiffs Stewart O'Nan, Abdi Nazemian, Brian Keene, Rebecca Makkai, and Jason Reynolds and

4   the proposed class (together "Plaintiffs") and Defendants Databricks, Inc. ("Databricks") and

5   Mosaic ML, LLC, formerly Mosaic ML, Inc. ("MosaicML") (together "Defendants") (collectively

6   the "parties"), have met and conferred and respectfully submit this Joint Case Management

7   Statement and Rule 26(f) Report in advance of the Initial Case Management Conference on June

8   21, 2024 in both cases. Pursuant to Federal Rule of Civil Procedure 26(f), conferences were held

9   on May 15, 2024 and May 21, 2024 between counsel for the *O'Nan* Plaintiffs and Defendants. On

10   June 6, 2024 counsel for all parties in both the *O'Nan* and *Makkai* Actions held a joint conference

11   and conferred concerning the topics set forth in Rule 26(f), the Northern District of California

12   Checklist for Rule 26(f) Meet and Confer Regarding Electronically Stored Information and

13   Guidelines for the Discovery of Electronically Stored Information, and the Standing Order for All

14   Judges of the Northern District of California—Contents of Joint Case Management Statement

15   (updated Nov. 30, 2023).

16   **1. Jurisdiction and Service**

17   This Court has subject matter jurisdiction over this action pursuant to 28 U.S.C. § 1331

18   (federal question jurisdiction), as this action alleges violations of the Copyright Act, 17 U.S.C. §

19   501. All parties have been served.

20   **2. Facts**

21   **Plaintiffs' Position**

22   The Complaints in the *O'Nan* Action and the *Makkai* Action each respectively allege one

23   count of direct copyright infringement against Defendant MosaicML and one count of vicarious

24   copyright infringement against Defendant Databricks. Plaintiffs are writers who own copyrights in

25   works allegedly used to train Defendants' large language models ("LLMs"). The Complaints each

26   _____

27   [1] The Court subsequently reset the due date of the Joint Case Management Statement to June 14, 2024 (*O'Nan* Dkt. 23).

28   [2] The Court subsequently reset the due date of the Joint Case Management Statement to June 14, 2024 (*Makkai* Dkt. 21).

1   respectively allege that MosaicML copied Plaintiffs' registered Works to train (i.e., the process by

2   which an Artificial Intelligence ("AI") program can "learn" how to anticipate and provide outputs

3   to prompts) its MPT language models in violation of the Copyright Act, 17 U.S.C. § 501. MosaicML

4   copied the Books3 dataset, which includes the *O'Nan* and *Makkai* Plaintiffs' Infringed Works, used

5   these copies to train MosaicML's LLMs, and made further copies to train additional models in the

6   MPT family. MosaicML's models—built on Plaintiffs' Infringed Works—directly compete with

7   Plaintiffs' Works. Defendant Databricks acquired MosaicML in July 2023. Databricks therefore has

8   the ability to control or curtail MosaicML's infringement of Plaintiffs' copyrights but has failed to

9   do so. Databricks has instead profited from MosaicML's infringement and is liable for vicarious

10   copyright infringement. The *O'Nan* and *Makkai* Plaintiffs have been injured by Defendants'

11   conduct.

12   **Defendants' Position**

13   MosaicML was founded in 2021 with a mission of making efficient training of machine

14   learning models accessible to enterprises. Databricks acquired MosaicML in July 2023. Founded

15   in 2013, Databricks seeks to democratize access to the tools necessary for all enterprises to benefit

16   from data analytics and AI.

17   Unlike providers of general-purpose generative AI models used by consumers, MosaicML

18   and Databricks focus on building tools that enable organizations to gain insights and make

19   predictions from their *own data*. From health care providers, who leverage Databricks' data

20   intelligence platform to deliver more personalized treatments for patients, to public sector agencies,

21   who use Databricks to help modernize their HR and hiring practices, Databricks provides tools that

22   organizations can use to best leverage their data.

23   Before its acquisition by Databricks, MosaicML developed and released large language

24   models called MosaicML Pretrained Transformer ("MPT") models. This litigation focuses on the

25   training of three of those MPT models. Training an LLM involves showing the model a large

26   quantity and wide diversity of textual data, which aids in teaching the model to understand how

27   words fit together syntactically and grammatically, as well as how words and sequences of words

28   work together to form concepts, meanings, and ideas. The purpose of these models is not to output

content it was trained on, but to create new material that never existed before, based on an understanding of human language and reasoning.

On March 8, 2024, Plaintiffs Stewart O'Nan, Abdi Nazemian, and Brian Keene filed a putative class action Complaint in the *O'Nan* Action (*O'Nan* Dkt. 1) against Defendants. The *O'Nan* Plaintiffs allege that MosaicML copied Plaintiffs' books from publicly available sources to train a series of MPT models. The Complaint alleges direct copyright infringement against MosaicML and vicarious copyright infringement against Databricks. On May 2, 2024, Defendants filed their Answer to the Complaint in the *O'Nan* Action (*O'Nan* Dkt. 39).

On May 2, 2024, Plaintiffs Rebecca Makkai and Jason Reynolds filed a putative class action Complaint in the *Makkai* Action (*Makkai* Dkt. 1) against Defendants. The *Makkai* Complaint is virtually identical to the *O'Nan* Complaint, asserts the same claims, and the *Makkai* Plaintiffs purport to represent the same class as the *O'Nan* Plaintiffs. On May 13, 2024, the Court entered an order relating the *O'Nan* Action and the *Makkai* Action (*O'Nan* Dkt. 45) (*Makkai* Dkt. 19). On May 29, 2024, Defendants filed their Answer to the Complaint in the *Makkai* Action (*Makkai* Dkt. 26).

Although discovery has just begun, Defendants preliminarily identify the following principal facts in dispute: (a) whether Defendants' use of Plaintiffs' asserted works to train the MPT models constituted fair use, including due to the purpose and transformative character of the use and the lack of an adverse effect of the use upon the potential market for or value of the allegedly copyrighted works; (b) whether each of the Plaintiffs have valid copyrights in each of the asserted works; (c) whether Plaintiffs have any contractual agreements affecting their alleged ownership of the asserted works; (d) whether and the extent to which Plaintiffs have enforced their alleged copyrights in each of the asserted works; (e) whether Defendants used each of Plaintiffs' asserted works; and (f) whether and the extent to which any Plaintiff has been injured by the alleged infringement.

1

3. **Legal Issues**

2

Plaintiffs' Position

3
      The legal and factual issue at the core of this case is: Whether Defendants' unlawful copying

4
has violated the Copyright Act, 17 U.S.C. § 501, *et seq.*

5
      Plaintiffs further believe there are numerous questions of law or fact common to the class,

6
and those issues predominate over any question affecting only individual class members. *See* Part

7
9, *infra*. Given the numerous common legal and factual issues, Plaintiffs do not believe that phased

8
discovery is necessary or appropriate in this case.

9
      Regarding Defendants' proposed sequencing of dispositive motions with respect to fair use

10
and class certification, Defendants' proposed schedule is in contravention of Rule 23. Plaintiffs, of

11
course, dispute that fair use is suitable for resolution via dispositive motion practice, and is instead

12
an issue that "must go to a jury." *Thomson Reuters Enter. Centre GmbH v. Ros Intelligence Inc.*, __

13
F. Supp. 3d __, 2023 WL 6210901, at *7 (D. Del. Sep. 25, 2023) (denying summary judgment on

14
fair use in generative AI case). In any event, any determination regarding fair use would be an issue

15
common to the class, and thus does not require determination prior to class certification. Further,

16
Defendants' proposal would deny Plaintiffs the opportunity to seek summary judgment before class

17
certification due to the one-way intervention rule.

18
      "[T]he history of the development of Rule 23(c)(2) makes clear that the rule was adopted to

19
prevent 'one-way intervention'—that is, the intervention of a plaintiff in a class action after an

20
adjudication favoring the class had taken place." *Schwarzschild v. Tse*, 69 F.3d 293, 297 (9th Cir.

21
1995). Under the one-way intervention rule, district courts generally do not grant motions for

22
summary judgment by plaintiffs on the merits of a class action until the class has been properly

23
certified and notified. *Moore v. Mars Petcare US, Inc.*, No. 16-CV-07001-MMC, 2021 WL

24
6063577, at *2 (N.D. Cal. Dec. 22, 2021). Nothing prevents Defendants from seeking early

25
summary judgment before class certification even if a schedule contemplates dispositive motions

26
after class certification. Such a motion, however, would waive the right to have any judgment bind

27
the proposed class. *Schwarzschild*, 69 F.3d at 297. Under Defendants' proposed schedule, however,

28
plaintiffs would be required to file summary judgment motions before class certification, in violation

JOINT CASE MANAGEMENT STATEMENT AND RULE 26(f) REPORT

1   of the one-way intervention rule. In other words, Defendant's proposal precludes Plaintiffs from

2   moving for summary judgment *after* class certification, which would run counter to the framework

3   provided by Rule 23. Though Defendants cite Judge Chhabria, other courts in this district have

4   disagreed with the approach—Judge Martínez-Olguín rejected a similar schedule seeking to

5   sequence dispositive motions before class certification. *Tremblay v. OpenAI Inc.*, Case No. 23-cv-

6   03223 (N.D. Cal.), ECF No. 77 at 12:8-13:9 (recognizing in generative AI class action that

7   sequencing summary judgment before class certification "will deny [Plaintiffs] the opportunity to

8   be able to seek summary judgment" and denying request to sequence summary judgment before

9   class certification).

10   **Defendants' Position**

11   Defendants deny any liability for the alleged unauthorized use of any of Plaintiffs' or

12   putative class members' registered copyrighted works during the training of MosaicML's MPT

13   models. To the extent MosaicML made any unauthorized copies of Plaintiffs' or putative class

14   members' registered copyrighted works, such copying constituted fair use under 17 U.S.C. § 107,

15   given, among other factors, the purpose and transformative character of the use and the effect of the

16   use upon the potential market for or value of the allegedly copyrighted works.

17   Defendants maintain that fair use can dispose of this case and should be resolved before

18   consideration of class certification. *See Authors Guild, Inc. v. Google Inc.* ("*Google Books I*"), 721

19   F.3d 132, 134-35 (2d Cir. 2013) (holding district court erred in ruling on class certification in a

20   putative class action copyright case involving the alleged infringement of books prior to resolving

21   defendant's fair use defense). Judge Chhabria took a similar approach in the pending *Kadrey v.*

22   *Meta* copyright case. *Kadrey et al v. Meta Platforms, Inc.*, Case No. 3:23-cv-03417-VC (N.D. Cal.

23   Jan. 12, 2024), Dkt. No. 85 (transcript of Jan. 12, 2024 proceedings) (ordering summary judgment

24   before class certification in a generative AI copyright class action on behalf of the same putative

25   class of authors, notwithstanding the "one-way intervention" issue Plaintiffs raise here).

26   Importantly, the fair use defense may moot class certification questions and greatly

27   streamline, if not dispose of, this litigation. *See Google Books I*, 721 F.3d at 134 ("resolution of

28   Google's fair use defense in the first instance will necessarily inform and perhaps moot our analysis

1 of many class certification issues" (collecting cases)). The Court in *Google Books I* also concluded

2 that holding class certification in abeyance until the fair use defense had been resolved would not

3 prejudice either party's interests. *Id.* at 135. Plaintiffs point out that the court in *Thomson Reuters*

4 denied summary judgment on fair use, but that court specifically did so based on the unique issues

5 and evidence in that case, not on any blanket rule that fair use questions must go to jury in every

6 case. *See Thomson Reuters Enter. Ctr. GmbH v. Ross Intelligence Inc.*, No. 1:20-CV-613-SB, 2023

7 WL 6210901, at *7-10 (D. Del. Sept. 25, 2023). Indeed, Google ultimately won the *Google Books*

8 case at summary judgment on the fair use defense. *Authors Guild v. Google, Inc.* ("*Google Books

9 II*"), 804 F.3d 202 (2d Cir. 2015) (affirming grant of summary judgment).

10   Defendants also preliminarily identify the following legal issues: (a) whether any use of

11 Plaintiffs' asserted works constituted fair use; (b) whether Plaintiffs' and putative class members'

12 alleged copyrights in their works are invalid or unenforceable, thereby causing them to lack standing

13 to assert their claims; (c) whether Plaintiffs' and putative class members' claims to copyrights in

14 their works are subject to an express or implied license, assignment, transfer, waiver, abandonment,

15 forfeiture, and/or doctrines of other equitable defenses; (d) whether Plaintiffs and putative class

16 members are entitled to injunctive relief; (e) whether Defendants' alleged conduct was innocent and

17 not willful; (f) whether Plaintiffs and putative class members suffered any injury as a result of the

18 alleged conduct; (g) whether Plaintiffs and putative class members failed to mitigate their alleged

19 damages, and (h) whether Plaintiffs can satisfy the requirements of class certification, including

20 commonality, typicality, adequacy of representation, and predominance of common questions of

21 law and fact.

22   **4. Motions**

23   There are no pending motions at this time.

24   As noted in Section 10 (Related Cases) below, the *O'Nan* and *Makkai* Actions have been

25 judicially related. Plaintiffs are discussing formal consolidation, and Defendants believe the cases

26 should be consolidated. If Plaintiffs do not move for the cases to be consolidated, Defendants intend

27 to move for the cases to be consolidated.

28

As noted in Section 8 (Discovery) below, Defendants do not seek to fully bifurcate discovery, but they propose a phased approach to discovery in which the first phase would focus primarily on the issue of fair use. Defendants intend to move for summary judgment on the issue of fair use, which would dispose of the case and avoid complex class certification proceedings and other merits issues. In accordance with the Court's Standing Order, which requires leave of the Court for a party to file more than one summary judgment motion, Defendants intend to request leave to move for summary judgment on the case-dispositive issue of fair use, while later moving for summary judgment on other issues if the case has not been dismissed. Plaintiffs believe that a phased approach to discovery is neither necessary nor appropriate in this matter.

The parties hope to resolve discovery disputes without judicial intervention but anticipate that discovery motions may be necessary.

5. Amendment of Pleadings

The parties do not anticipate any substantive amendments to the pleadings at this time.

6. Evidence Preservation

The parties certify that they have reviewed this Court's Guidelines Relating to the Discovery of Electronically Stored Information and have met and conferred pursuant to Rule 26(f) regarding reasonable and proportionate steps taken to preserve evidence relevant to the issues reasonably evident in the case. The parties are aware of their obligations and have taken reasonable steps to preserve potentially relevant evidence. The parties will meet and confer concerning ESI.

7. Disclosures

The parties will serve Federal Rule of Civil Procedure 26(a) initial disclosures by June 20, 2024.[3]

8. Discovery

The *O'Nan* Plaintiffs had an initial Rule 26(f) meet and confer with Defendants on May 15, 2024, and a second meet and confer to discuss an ESI Protocol and discovery stipulations on May

---

[3] The parties have agreed to exchange initial disclosures two weeks from June 6, 2024, the date of the joint Rule 26(f) conference attended by counsel for Defendants and both the *O'Nan* and *Makkai* Plaintiffs.

21, 2024. After these calls, the *O'Nan* and *Makkai* Plaintiffs agreed to coordinate all discovery, including discovery stipulations. The *O'Nan* Plaintiffs served written discovery on Databricks on May 29, 2024, which the *Makkai* Plaintiffs join. On June 6, 2021, counsel for all parties met and conferred on a coordinated discovery plan, addressed Rule 26(f) requirements, and continued negotiations on discovery stipulations.

a. Scope of Discovery

The parties agree that discovery is governed by the Federal Rules of Civil Procedure. The parties will meet and confer on any discovery disputes that may arise in an effort to resolve them without court intervention. The parties agree to electronic service.

Plaintiffs' Statement: For the reasons stated in Part 3, *supra*, Plaintiffs do not believe that phased discovery is necessary or appropriate in this matter. Plaintiffs anticipate conducting discovery, including documents and testimony from Databricks and Mosaic regarding, among other things: the development and training of Defendants' MPT models; the training data for the MPT models; Defendants' creation and/or copying of the Books3 and "RedPajama—Books" dataset, and general awareness of, and policies about, using copyrighted materials and/or pirated libraries in training data; sources of funding for the development of the MPT models; financial market and profitability of the MPT models; licensing deals and negotiations for training data; knowledge of, and participation within, the AI training data market. Plaintiffs reserve all rights to update, add, or amend this list as additional information becomes available.

Plaintiffs are opposed to phasing or bifurcating discovery in any way. The fair use defense must be decided on a full discovery record. *Thomson Reuters*, 2023 WL 6210901, at *7. Imposing topical limitations on discovery empowers Defendants to narrow their responses to discovery requests based on internal determinations about relevance only as to their defense. Because fair use is an affirmative defense to the underlying claims of infringement and is a common issue to all class members, full discovery on the merits of Defendants' infringement is also relevant to the resolution of the fair use defense. Accordingly, any phasing of discovery would only frustrate the efficient resolution of this litigation and no limitations should be placed on discovery the parties may propound based on the sequencing of dispositive motions.

*Subjects of Discovery from Defendants***:** The following is a non-exhaustive list of subjects that Plaintiffs have sought or plan to seek discovery on. Plaintiffs reserve all rights to expand upon these subject areas as the case proceeds:

- The interactions between class members (including Plaintiffs) and Defendants, including but not limited to requests and/or demands from the owners of works used to train Defendants' language models regarding Defendants' use of those works.

- Defendants' conduct that caused the language models to ingest and distribute Plaintiffs' and the class's works, including but not limited to Defendants' knowledge regarding those actions and their consequences.

- The planning for, creation, operation, modification, distribution and maintenance of Defendants' language models, including but not limited to source code.

- Defendants' efforts to obtain licenses or other permission to use works used to train their language models.

- Defendants' organizational structures and corporate documents.

- Contracts, licenses or other agreements between Defendants, Defendants and Plaintiffs, and Defendants and third parties that relate to Defendants' language models.

- Information regarding all materials used to train Defendants' language models, including the materials themselves.

- Facts related to affirmative defenses raised by Defendants.

- Defendants' policies and practices regarding any usage of copyrighted material.

- Defendants' interactions with government or regulatory entities regarding Defendants' language models, including document submissions, requests for information or documents, testimony, and correspondence.

- Defendants' participation and/or involvement in other lawsuits, private administrative proceedings, or regulatory proceedings regarding generative AI language models.

Defendants' Statement: Defendants propose a phased approach to discovery. The first phase would focus primarily on discovery pertaining to Defendants' affirmative defense of fair use. At the conclusion of the first phase, Defendants anticipate dispositive briefing on the issue of fair use. Should any claims remain after dispositive briefing on the issue of fair use, the second phase of discovery would pertain to remaining merits claims and class certification.

Subjects of Discovery from Plaintiffs: Defendants anticipate conducting discovery, including documents and testimony from Plaintiffs regarding, among other things, the following topics. Defendants reserve all rights to update, add, or amend this list as additional information becomes available:

- The alleged copyrights for Plaintiffs' works.

- Plaintiffs' efforts to enforce the alleged copyrights for their works.

- The market for Plaintiffs' works.

- The marketing and licensing of Plaintiffs' works

- Use of Plaintiffs' works in connection with machine learning or models using artificial intelligence.

b. Completion of Discovery

The Court has not set a deadline for the close of fact discovery. Plaintiffs propose that discovery should be completed by July 28, 2025. Defendants propose that fact discovery related primarily to the fair use defense be completed by April 18, 2025, expert discovery related primarily to the fair use defense be completed by July 25, 2025, and that a further Case Management Conference be conducted after ruling on summary judgment on the fair use issue, if any claims remain.

c. Discovery Limitations

The parties have met and conferred regarding the discovery limits set forth in the Federal Rules of Civil Procedure.

Given the nature of the legal and factual issues in this case, Plaintiffs believe that expansion of the discovery limits as set forth in the Federal Rules of Civil Procedure will be necessary in this case. Plaintiffs have conferred with Defendants regarding the likelihood they will

1 require expansion of the number of depositions and interrogatories and believe it is premature to

2 limit the number of depositions and interrogatories.

3   Defendants maintain that any expansion of the default rules is unwarranted and not

4 proportional to the needs of the case, particularly because MosaicML was a small startup company

5 with well under 100 employees at the time of the events in question. Plaintiffs have provided no

6 justification for their proposed expansion.

7   Plaintiffs seek to have a protocol established for depositions, and Defendants believe such a

8 protocol is unnecessary.

9   Plaintiffs also seek an expert discovery order. Defendants believe that such an order is

10 unnecessary but are prepared to continue conferring with Plaintiffs' counsel about this subject.

11 In accordance with Judge Breyer's Standing Order for Civil and Criminal Cases, the parties will

12 seek a referral to the Magistrate Judge to resolve any potential disputes regarding deposition and

13 expert discovery protocols.

d. Preservation and Production of Electronically Stored Information

15   The parties have reviewed the Guidelines Relating to the Discovery of Electronically

16 Stored Information ("ESI Guidelines"). The parties have and are continuing to meet and confer

17 regarding a joint stipulated ESI Protocol. Defendants seek an ESI Protocol largely based on this

18 Court's model, while Plaintiffs seek a more detailed ESI Protocol which reflects contemporary

19 ESI best practices established by practitioners and jurists since the promulgation of the model

20 order. The parties will continue to meet and confer on this topic in an attempt to come to

21 agreement on an ESI Protocol to submit to the Court.

e. Protective Order

23   The parties have begun to discuss an appropriate Protective Order for confidential and more

24 sensitive technical materials based on the Court's models, and the parties are working on a joint

25 stipulated Protective Order to be entered by the Court. The parties will continue to meet and confer

26 on this topic in an effort to come to an agreement on a Protective Order to submit to the Court.

27

28

f. Privilege

The parties have and are continuing to meet and confer regarding a protocol for privilege issues, including privilege logging, which will be incorporated into the joint stipulated ESI Protocol.

With respect to the 502(d) Agreement, the parties have and are continuing to meet and confer, and have agreed to address a Rule 502(d) order as part of the Protective Order, noted above.

g. Deposition Logistics

The parties will meet and confer in good faith on a deposition-by-deposition basis to determine the appropriateness and feasibility regarding in-person and remote depositions. The parties expect that any in-person depositions will be conducted in the Northern District of California or elsewhere in the United States. The parties are unaware of depositions to be taken outside the United States at this time. The parties agree to meet and confer should any disputes arise as to location of depositions and/or use of remote depositions.

Plaintiffs believe that a deposition protocol is appropriate to set mutually agreed upon guidelines for depositions in this case.

h. Discovery Issues and Potential Disputes

The parties are prepared to commence discovery. The parties have not identified any specific discovery dispute and are prepared to meet and confer on any other discovery dispute that may arise. In accordance with Judge Breyer's Standing Order for Civil and Criminal Cases, the parties will seek a referral to the Magistrate Judge to resolve any potential disputes regarding discovery issues.

i. Narrowing of Issues

No issues have yet been narrowed by agreement or by motion. The parties are prepared to meet and confer about narrowing potential issues should the circumstances of the case change.

9. Class Action

This case is a putative class action. Counsel for both sides have reviewed the Procedural Guidance for Class Action Settlements.

Plaintiffs' Statement: Plaintiffs believe that the class is so numerous that joinder of all members is impracticable. Plaintiffs do not know the exact size of the class since that information is within the exclusive control of Defendants. Upon information and belief, Plaintiffs believe and

1   allege that the number of class members is in the millions. Plaintiffs allege that, like all class

2   members, their copyright-protected works were used to train Defendants' language models.

3   Plaintiffs allege all members of the class have been similarly injured by Defendants' conduct.

4   Plaintiffs believe there are numerous questions of law or fact common to the class, and those issues

5   predominate over any question affecting only individual class members. These common legal and

6   factual issues include the following:

- Whether Defendants violated the copyrights of Plaintiffs and the class when they downloaded copies of Plaintiffs' copyrighted works and used them to train Defendants' language models.

- Whether this Court should enjoin Defendants from engaging in the unlawful conduct alleged herein, and the appropriate scope of that injunction.

- Whether any affirmative defense (including fair use) excuses Defendants' conduct.

- The amount of damages.

14   Plaintiffs do not believe that phased discovery is necessary or appropriate in this case.

15   **Defendants' Statement:** Defendants dispute Plaintiffs' class certification arguments and

16   maintain that the proposed putative class cannot be certified under the requirements of Rule 23.

17   Defendants also believe that consideration of class certification issues would be inappropriate and

18   inefficient until after the Court rules on an early dispositive motion on Defendants' affirmative

19   defense of fair use. *See Google Books I.*, 721 F.3d at 134-35 (holding district court erred in ruling

20   on class certification in a class action copyright case involving the alleged infringement of books

21   prior to resolving defendant's fair use defense). As noted above, Judge Chhabria took a similar

22   approach in *Kadrey v. Meta*. *Kadrey et al. v. Meta Platforms, Inc.*, Case No. 3:23-cv-03417-VC

23   (N.D. Cal. Jan. 12, 2024), Dkt. No. 85 (transcript of Jan. 12, 2024 proceedings) (ordering summary

24   judgment before class certification in a generative AI copyright class action on behalf of the same

25   putative class of authors).

26   Accordingly, as described in Section 8 (Discovery), Defendants believe that a phased

27   approach to discovery is most efficient in this case. The first phase would focus primarily on

28   discovery pertaining to Defendants' affirmative defense of fair use. At the conclusion of the first

phase, Defendants anticipate dispositive briefing on the issue of fair use. Should any claims remain after dispositive briefing on the issue of fair use, the second phase of discovery would pertain to any remaining merits claims and class certification.

If the case survives summary judgment on the fair use defense, Defendants do not believe at this time that class certification will be appropriate. Among other factors, Defendants expect that individual questions of law and fact will predominate over common questions, including whether class members hold valid copyright registrations, the existence and timing of registration of the relevant works, whether the works were created as works for hire, whether putative rights-holders have assigned or transferred their rights, the nature of and market for each work, the marketing and licensing history of each work, and whether each author or copyright owner has authorized or objects to use of the work in connection with machine learning or models using artificial intelligence.

10. Related Cases

Pursuant to the Court's May 13, 2024 Order (*O'Nan* Dkt. 45) (*Makkai* Dkt. 19), the later-filed case *Makkai et al. v. Databricks, Inc. et al.*, No. 3:24-cv-02653 (N.D. Cal.) has been related to *O'Nan et al. v. Databricks, Inc. et al.*, No. 3:24-cv-01451 (N.D. Cal.). Counsel are discussing consolidation of these two cases.

The parties are unaware of any other cases brought on behalf of authors with registered copyrights involving Defendants' large language models other than the two cases involved here.

11. Relief

Plaintiffs' Position: The *O'Nan* and *Makkai* Plaintiffs seek statutory and other damages under 17 U.S.C. § 504 for Defendants' violations of the copyrights of Plaintiffs and the class, including an award of reasonable attorneys' fees under 17 U.S.C. § 505 or other applicable statutes. Plaintiffs further seek destruction or other reasonable disposition of all copies Defendants made or used in violation of the exclusive rights of Plaintiffs and the class, pursuant to 17 U.S.C. § 503(b). Plaintiffs seek pre-and post-judgment interest on damages awarded to Plaintiffs and the class, at the

1   highest legal rate from and after March 8, 2024, the date this class action Complaint was first served

2   on Defendants.

3   **Defendants' Position:** Defendants seek a judgment in Defendants' favor denying Plaintiffs'

4   requested relief in the *O'Nan* and *Makkai* Complaints and dismissing the Complaints with prejudice.

5   Should Defendants prevail in their defenses, Defendants seek an award of their attorneys' fees, costs,

6   and expenses of litigation.

7   Defendants deny that Plaintiffs or any member of the putative class are entitled to any

8   damages, statutory or otherwise; any injunction or other equitable relief; any fees or costs; any pre-

9   or post-judgment interest; or any other relief of any kind. Defendants deny that any destruction or

10   other disposition of copies of Plaintiffs' works allegedly made by Defendants is warranted because

11   any such use constituted fair use.

12   **12. Settlement and ADR**

13   ADR Certifications pursuant to ADR Local R. 3-5(b) were filed by the *O'Nan* Plaintiffs on

14   May 9, 2024, (*O'Nan* Dkts. 41–43), and by Defendants in the *O'Nan* Action on May 16, 2024

15   (*O'Nan* Dkt. 46).

16   ADR Certifications pursuant to ADR Local R. 3-5(b) were filed by Plaintiff Rebecca Makkai

17   in the *Makkai* Action on June 14, 2024, (*Makkai* Dkt. 43)[4], and by Defendants on June 4, 2024

18   (*Makkai* Dkt. 36). The parties continue to meet and confer on an ADR plan.

19   **13. Other References**

20   The parties agree that this case is not suitable for binding arbitration, a special master, or the

21   Judicial Panel on Multidistrict Litigation at this time.

22   The parties do not oppose having discovery disputes heard by a magistrate judge if the Court

23   is inclined to appoint a discovery magistrate.

24   **14. Narrowing of Issues**

25   No issues have yet been narrowed by agreement or by motion.

26

27

28

---

[4] Plaintiff Jason Reynolds's ADR Certification is forthcoming.

1   As described in Section 8 (Discovery), Defendants believe that a phased approach to

2 discovery and early summary judgment will appropriately narrow the issues in the case. Plaintiffs

3 disagree that phased discovery is necessary or appropriate in this case.

4   **15. Scheduling**

5   **a. Expedited Trial Procedures**

6   The parties agree that these Actions are inappropriate for the Expedited Trial Procedure of

7 General Order 64.

8   **b. Case Schedule**

9   **Plaintiffs' Position:** Plaintiffs propose the following schedule[5]:

Event	Plaintiffs' Proposed Date
Parties must submit stipulated ESI Protocol or file 10-page joint letter brief. The parties shall simultaneously exchange sections of the joint letter brief in two phases with agreed-upon deadlines prior to the submitting to the Court.	June 17, 2024
Parties must submit stipulated Protective Order or file 10-page joint letter brief. The parties shall simultaneously exchange sections of the joint letter brief in two phases with agreed-upon deadlines prior to the submitting to the Court.	June 17, 2024
Rule 26(a)(1) initial disclosures due, including Supplemental Information Disclosures noted in the Rule 26(f) Report	June 20, 2024
Producing parties propose document custodians, non-custodial document sources, search terms, any agreed-upon or required custodian cellphone information	July 1, 2024
Requesting parties propose additional custodians and non-custodian document sources, and any additional search terms ("Requesting Party's Proposal")	July 10, 2024
Defendants shall each produce a sample production of data from each of their relevant Structured Data databases and model data	July 17, 2024

---

[5] On June 7, 2024, Plaintiffs have also indicated to Defendants that they intend to propose that the parties make Supplemental Information Disclosures. Defendants do not agree to Plaintiffs' proposed Supplemental Information Disclosures as they are beyond any obligations for disclosure required by Rule 26.

Case No. 3:24-cv-01451-CRB
Case No. 3:24-cv-02653-CRB

JOINT CASE MANAGEMENT STATEMENT AND RULE 26(f) REPORT

Event	Plaintiffs' Proposed Date
Date by which parties must submit an agreed upon deposition protocol or a joint letter brief regarding any disputed areas	July 17, 2024
Producing party to accept or reject the Requesting Party's Proposal	July 24, 2024
Parties shall complete their meet and confer process regarding Defendants' Structured Data sample, including raising any questions about the sample productions	July 24, 2024
Date by which parties submit joint letter brief to the Court regarding areas of dispute on custodial and non-custodial document sources (if any) and disputed search terms (if any)	July 31, 2024
Deadline for Plaintiffs to identify three priority custodians per Defendant	July 31, 2024
Defendants shall answer questions about the form and contents of their Structured Data and any perceived deficiencies therein in order to facilitate the timely production of an agreed upon Structured Data by August 29, 2024	August 7, 2024
Date by which parties must submit an agreed upon protocol for the authentication and qualification of documents as business records or a joint letter brief regarding any disputed areas.	August 22, 2024
Deadline for substantial completion of priority custodians' documents and production of structured data	September 12, 2024
Deadline for requesting parties' second phase search term proposal	October 10, 2024
Deadline for resolution of second phase search term proposals or filing of joint letter brief regarding any disputed issues	October 17, 2024
Substantial completion of production of documents and unstructured data in response to discovery requests served prior to July 14, 2024	November 14, 2024
Deadline for Plaintiffs' to file an amended pleading	30 days after substantial completion of document production

Event	Plaintiffs' Proposed Date
Producing party to produce privilege logs for any documents withheld on the basis of any privilege to date	November 21, 2024
Producing parties' file certifications of substantial completion of document productions and production of privilege logs, specifying any categories of requested documents not yet produced to the requesting party	November 21, 2024
Filing of Plaintiffs' Expert Report(s)	May 1, 2025
Filing of Defendants' Expert Report(s)	June 16, 2025
Close of Fact Discovery	July 28, 2025
Filing of Plaintiffs' Expert Rebuttal Reports	August 1, 2025
Filing of Plaintiffs' Motion for Class Certification	August 15, 2025
Filing of Defendants' *Daubert* Motions regarding Plaintiffs' Expert(s)	August 22, 2025
Filing of Defendants' Opposition to Motion for Class Certification	September 12, 2025
Filing of Plaintiffs' Opposition to Defendants' *Daubert* Motions Filing of Plaintiffs' *Daubert* Motions on Defendants' Expert(s)	September 17, 2025
Filing of Plaintiffs' Reply in Support of Motion for Class Certification	October 10, 2025
Filing of Defendants' Reply in Support of *Daubert* Motion(s) Filing of Defendants' Opposition to Plaintiffs' *Daubert* Motion(s)	October 17, 2025
Filing of Plaintiffs' Reply in Support of Plaintiffs' *Daubert* Motion	October 31, 2025
Hearing on Class Certification and *Daubert* Motions	*At the Court's convenience*

Event	Plaintiffs' Proposed Date
Deadline for filing Rule 56 Motions	January 22, 2026
Deadline for filing Oppositions to Rule 56 Motions	March 23, 2026
Deadline for filing Replies in Support of Rule 56 Motions	April 22, 2026
Hearing on Rule 56 Motions	*At the Court's convenience*
Trial Ready Date	August 3, 2026

Defendants' Position: Defendants believe that Plaintiffs' proposed schedule above is unnecessarily complicated, detailed, and rigid, and will only invite further disputes. As described in Section 9 (Class Action), Defendants believe that consideration of class certification issues would be inappropriate and inefficient until after the Court rules on an early dispositive motion on Defendants' affirmative defense of fair use.

Accordingly, Defendants respectfully propose a schedule for fact and expert discovery focused primarily on the issue of fair use, followed by dispositive motions which focus on the fair use defense. After the Court's ruling on summary judgment on the issue of fair use, if any issues remain, Defendants propose further scheduling of proceedings on class certification and determination of a trial date, if necessary, through further Case Management Conferences:

Event	Defendants' Proposed Date
Close of Fact Discovery on Plaintiffs' copyright claim/fair use	April 18, 2025
Deadline for Opening Expert Reports on Plaintiffs' copyright claim/fair use	May 16, 2025
Deadline for Rebuttal Expert Reports	June 20, 2025
Close of Expert Discovery	July 25, 2025

Event	Defendants' Proposed Date
Deadline for Summary Judgment Motions on Plaintiffs' copyright claim/fair use [6]	September 5, 2025
Deadline for Oppositions to Summary Judgment Motions	October 10, 2025
Deadline for Replies to Summary Judgment Motions	November 7, 2025
Hearing on Summary Judgment Motions	*To be set by the Court*

16. Trial

The case will be tried before a jury. Plaintiffs anticipate a 14-day trial. Defendants anticipate a 7-day trial.

17. Disclosure of Non-Party Interested Entities or Persons

The *O'Nan* Plaintiffs filed a Certification of Interested Entities or Persons along with the Complaint on March 8, 2024 (*O'Nan* Dkt. 5). The *Makkai* Plaintiffs filed their Certification of Interested Entities or Persons on June 14, 2024 (*Makkai* Dkt. 44). Defendants filed a Certification of Interested Entities or Persons on March 25, 2024, in the *O'Nan* Action (*O'Nan* Dkt. 30) and on May 29, 2024 in the *Makkai* Action (*Makkai* Dkt. 25), certifying that there is no conflict or interest (other than the named parties) to report.

18. Professional Conduct

All attorneys of record for the parties have reviewed the Guidelines for Professional Conduct for the Northern District of California.

19. Other Matters

The parties are not aware of other matters that may facilitate the resolution of these cases.

---

[6] In accordance with the Court's Standing Order, which requires leave of the Court for a party to file more than one summary judgment motion, Defendants intend to request leave to move for summary judgment on the case-dispositive issue of fair use, while later moving for summary judgment on other issues if the case has not been dismissed.

1   DATED: June 14, 2024

2

3 By: */s/ Joseph R. Saveri*   By: */s/ Jedediah Wakefield*
   Joseph R. Saveri (State Bar No. 130064)   Jedediah Wakefield
4 Cadio Zirpoli (State Bar No. 179108)   Ryan Kwock
   Christopher K.L. Young (State Bar No.   **FENWICK & WEST LLP**
5 318371)   555 California Street, 12th Floor
   Elissa Buchanan (State Bar No. 249996)   San Francisco, CA 94104
6 Evan Creutz (State Bar No. 349728)   Telephone: (415) 875-2300
   JOSEPH SAVERI LAW FIRM, LLP   Email: jwakefield@fenwick.com
7 601 California Street, Suite 1505   rkwock@fenwick.com
   San Francisco, CA 94108
8 Telephone: (415) 500-6800   David Lloyd Hayes
   Facsimile: (415) 395-9940   **FENWICK & WEST LLP**
9 Email: jsaveri@saverilawfirm.com   801 California Street
10   czirpoli@saverilawfirm.com   Mountain View, CA 94041
               cyoung@saverilawfirm.com   Telephone: (650) 988-8500
11   eabuchanan@ saverilawfirm.com   Email: dhayes@fenwick.com
               ecreutz@saverilawfirm.com
12   Charles Moulins (admitted *pro hac vice*)
13 Matthew Butterick (State Bar No. 250953)   **FENWICK & WEST LLP**
   1920 Hillhurst Avenue, #406   902 Broadway, Suite 14
14 Los Angeles, CA 90027   New York, NY 10010
   Telephone: (323) 968-2632   Telephone: (212) 430-2600
15 Facsimile: (415) 395-9940   Email: cmoulins@fenwick.com
   Email: mb@buttericklaw.com
16   Deena J.G. Feit (admitted *pro hac vice*)
17 Brian D. Clark (admitted *pro hac vice*)   **FENWICK & WEST LLP**
   Laura M. Matson (admitted *pro hac vice*)   401 Union Street, 5th Floor
18 Arielle S. Wagner (admitted *pro hac vice*)   Seattle, WA 98101
   Eura Chang (admitted *pro hac vice*)   Telephone: (206) 389-4510
19 **LOCKRIDGE GRINDAL NAUEN PLLP**   Email: dfeit@fenwick.com
   100 Washington Avenue South, Suite 2200
20 Minneapolis, MN 55401   *Counsel for Defendants*
   Telephone: (612) 339-6900   *Databricks, Inc. and Mosaic ML, LLC,*
21 Facsimile: (612) 339-0981   *formerly MosaicML, Inc.*
22 Email: bdclark@locklaw.com
               lmmatson@locklaw.com
23   aswagner@locklaw.com
               echang@locklaw.com
24

25 *Counsel for Individual and Representative*
   Plaintiffs and the Proposed Class in the
26 O'Nan *Action*

27

28

1   Bryan L. Clobes (*pro hac vice*)
     Alexander J. Sweatman (*pro hac vice*)
2   Mohammed R. Rathur (*pro hac vice*)
     CAFFERTY CLOBES MERIWETHER
3   **& SPRENGEL LLP**
     135 South LaSalle Street, Suite 3210
4   Chicago, IL 60603
     Telephone: 312-782-4880
5   bclobes@caffertyclobes.com
     asweatman@caffertyclobes.com
6   mrathur@caffertyclobes.com

7   Amy E. Keller (*pro hac vice*)
     Nada Djordjevic (*pro hac vice*)
8   James A. Ulwick (*pro hac vice*)
9   **DiCELLO LEVITT LLP**
     Ten North Dearborn Street, Sixth Floor
10 Chicago, Illinois 60602
     Tel. (312) 214-7900
11   akeller@dicellolevitt.com
         ndjordjevic@dicellolevitt.com
12   julwick@dicellolevitt.com
13
     David A. Straite (*pro hac vice*)
14 **DICELLO LEVITT LLP**
     485 Lexington Avenue, Suite 1001
15 New York, NY 10017
     Tel. (646) 933-1000
16   dstraite@dicellolevitt.com
17
18 *Attorneys for Plaintiffs and the Proposed*
     Class in the Makkai *Action*
19
20
21
22
23
24
25
26
27
28

1

ATTESTATION OF PURSUANT TO CIVIL L.R. 5-1(i)(3)

2   This document is being filed through the Electronic Case Filing (ECF) system by attorney

3 Joseph R. Saveri. By their signature, Joseph R. Saveri attests that he has obtained concurrence in

4 the filing of this document from each of the attorneys identified on the caption page and in the

5 above signature block.

6   Dated: June 14, 2024   By */s/ Joseph R. Saveri*

7   Joseph R. Saveri

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28